

# 他者の認識の推定に基づく知的インタラクションの試み

## Effect of estimating other's recognition in intelligent human interaction

大森隆司<sup>\*1</sup>  
Takashi Omori

奥谷一陽<sup>\*2</sup>  
Kazuhiro Okutani

<sup>\*1</sup> 玉川大学工学部  
School of Engineering, Tamagawa University

<sup>\*2</sup> 日本ケミコン株式会社  
Nippon Chemi-Con Corp.

In a communication scene between human, we estimate the belief of other person on outer world, and change own action based on the estimation. Belief inference is indispensable for the human-human, and for the human-agent interaction. However, there are few models that represent dynamics of belief in the interaction scene. So, in this study, we use a chasing game that require the belief process for realizing higher level interaction, and evaluate a belief model we developed by a computer simulation. We also discuss on a possibility of cheating that makes use of the belief estimation process.

### 1. はじめに

コミュニケーションには、言語・身振り・表情のように外部からの直接観測が可能な過程と、意図や信念のように外部からの直接観測はできない心的な過程が存在する。前者は直接観測による情報伝達により実現され、後者は前者の観測からはじまる推論によって相手の心的状態を推定することで、自己の心的状態や判断の変化を引き起こすことで実現される。

現在、コミュニケーションの手段として多くの機器が開発されているが、その多くはユーザーの観測可能な部分についてのインターフェース特性の理解とそれに基づく機器の機能設計に基づいている[菅野 2006]。しかし、コミュニケーションは相手との相互理解により相互の行動を適切に変えていくための手段であり、その本質はむしろ直接には観測できない心的相互作用にある。さまざまな状況で相手の行動の意図を汲んで適切に応えるためには、相手の表面的な行動の解釈に加え、直接観測はできない相手の信念や意図を推定しての意味解釈が必要である[大沢 1989]。本稿ではこのような心的過程も含めた相互作用を知的インタラクションと呼ぶ。

これまで、心的過程を記述する手法として論理を使う方法が研究されてきた[大沢 1989][菅野 2006]。しかし現状では、それらは思考の論理的記述に終わっており、現実的な場面での運用には至っていない。しかし意図や信念はコミュニケーションの重要な要因であり、これを無視することはできない。また横山は、意図を人間の行動決定に影響する隠れ変数として定式化し、ハンターゲームという協調ゲームの計算機シミュレーションによりその意図推定の方式とその運用アルゴリズムを定式化し、その有用性を示している[横山 2009]。また阿部はそのモデルをロボットと子供の遊び場面に適用し、子供の興味状態の推定による同調性が相互作用に有用であると提唱している[阿部 2011]。しかし、同様に重要と考えられる信念については、いまだ理論の範囲にとどまっており、現実的な場面での運用についてのモデルの検討は少ない。そこで本研究では信念に焦点を当て、知的インタラクション場面におけるそのモデルと運用について検討する。

本稿でいう信念とは、行動主体が考えるところの外部世界についての事実(すなわち外界についての認識)である。信念は

個体が環境を観測することで獲得され、環境の変化や行動主体の内部でアップデートされるが、信念の表現(すなわち記憶)が時間とともに減衰することで消えていく。その間、行動主体は環境についての自身の信念に基づいて行動を決める。よって行動決定過程と信念は不可分の関係にある。

一方で知的インタラクション場面では、我々は他者の行動を観測してそこから対象のもつ信念を推定する。他者の信念の推定はその行動の予測に直結する。他者の行動の予測を織り込んで自己の行動を決める方略は、現実の予測という問題と自身の行動決定という問題を分離させることで、行動主体に求められる情報処理を単純化する。このように他者の信念の推定は、知的インタラクションの基本的かつ重要な問題である。

そこで本研究は、他者の信念の推定と行動予測に関わる問題点を整理するため、他者の信念の推定が必要とされる行動決定ゲームを開発する。そして、その世界での行動決定モデルの構築とシミュレーションによる検証を通じて、他者と自己の信念を通じた相互作用の基本的な原理を検討する。

### 2. 2. 信念に関わる現象のモデル化の試み

#### 2.1 見えない他者についての信念

本稿では信念を「行動主体が考えるところの外部世界についての事実」とする。信念は、外界の観測およびそこから推論によって獲得され、状況が変化するとその認識に基づき更新される。そのため信念は、各行動主体が個別に持ち、観察がなくとも脳内の短期・長期の記憶を通じて継続される。

本稿の論点は、信念は主体の行動決定に影響するため、逆に特定の状況では他者の行動からその信念(外界についての認識状態)を推定できるであろうし、それを推定することが自身にとって有利であろうという点にある。では実際、どういう情報から信念は推定できるのか、あるいはどういう場面でその推定が自身にとって有利なのか、といったことをモデル的に検討したい。

その目的のため本研究では、「おにごっこ」などの追跡ゲームにおいて他者が見えなくなった時の、その対象の動きや位置に関する信念を扱う。おにごっこで対象(オニまたは子)が見えなくなった時、我々はその対象についての知識を用いてその現在状態を推定する。例えば我々は、遮蔽物により見えなくとも他者は継続的に存在すると考える。そして相手が見えなくなっても、オニは普通は追ってくるし、子はオニから逃げる行動が一般に予想される。

しかしこのような他者は、それ自身が自ら考えて行動する行動主体であり、その行動予測は決まったパターンに従うモデルだけでは困難と考えられる。例えば、オニはただ子を追いかけるだけでなく、後述のように時として逆回りをして子を待ちうける、という戦略もとりうる。その場合には、戦略レベルでの行動選択の議論をしなければならない。

ここで必要になるのは、他の行動主体の行動選択についての信念である。「おにごっこ」では、相手が見えている間は相手を観測してその現在状態から最適な行動をとることができる。しかし相手が見えない場合には、見えない間に相手がどのような行動戦略を取るかを考え、その相手の戦略に基づいた相手の行動予測を行わねばならない。すなわち、相手が見えなくなる「おにごっこ」では、他者の行動戦略の選択の推定もまた、個々の戦略における行動予測と同様に重要である。すなわち、「遮蔽物のあるおにごっこ」では、他者について取りうる戦略、その戦略に従う場合の行動予測、状況に応じた戦略の選択をモデル化しなければならない、ということになる。

ここで問題を複雑にするのは、推定される他者もまた自身の行動や相手についての信念を持ち、それが行動に影響を及ぼすことである。これまで横山はそのような場面での行動決定や他者の意図の誘導のモデルについて論じてきているが[横山 2009]、その推定の階層は浅く、他者の信念の推定までは及んでいない。例えば、行動主体 A,B が存在したとする。A はそれまで観測した B の行動と外界状況から、B の信念状態を推定し、その信念と B についてのモデルから B のこれからの行動を予測する。A はその予測、すなわち自己の認識に基づき行動するので、これは A が B について推定した事柄を信念として持つということになる。この A の B についての推定には、B が何を認識しているか、すなわち B の信念も含まれている。仮に、その推定された B の信念の中に A についての認識が含まれているなら、相互推定問題は信念にまで及ぶことになる。しかし本稿では、信念に関わるモデル化として問題を複雑化させることは避けた、相互推定問題は議論しない。

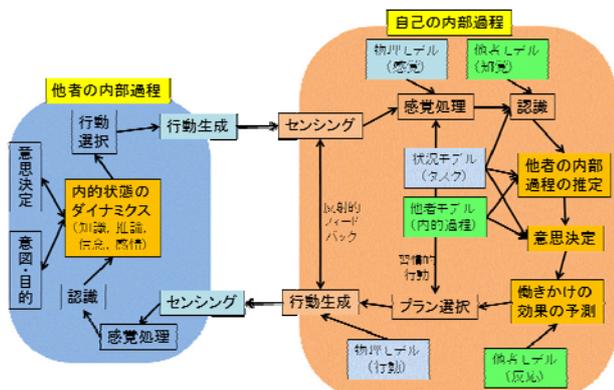


図1 信念や他者意図推定を含む行動決定過程の全体像

## 2.2 他者の行動戦略についての信念：一貫性

行動主体は、他者の行動についての知識、すなわち他者モデルを持つことで、より効果的な追跡(逃走)行動が選択できる。その際、主体は他者を取りうるいくつかの戦略に応じたモデル群のうち一つを選択することになる。そのモデルは相手が見えている間に相手を観測して学習・選択し、見えない時にそれを運用することになるが、他者の戦略モデルの推定自体には、相手が見えなくなった後もそれに従うという認識、すなわち相手の行動決定過程の安定性についての仮定が含まれている。

そこで本研究では、「おにごっこ」における行動決定の戦略を表す他者モデルのパラメータとして、行動の一貫性を考える。「一貫性」とは、行動主体が目的に対して無駄なく継続して行動することの指標である。一貫性の高い行動とは、目的地に向かって寄り道せず最短経路で向かうことであり、低い行動とはゆらぎを多く含んだ行動となる。それを表すパラメータは、対象を継続して観測することで推定される。「おにごっこ」における信念の立場からは、その行動主体が持つ認識の正しさについての確信の程度となり、他者の信念推定としてはその行動戦略がしばらくは継続されるであろうという予測につながる。

## 2.3 「おにごっこ」ゲームとゲーム世界

本研究では、信念の特性を検討してモデル構築につなぐために、遮蔽物のある場での追跡ゲーム「おにごっこ」を考える。「おにごっこ」では、遮蔽となる障害物が無い場合には、追跡する「オニ」と逃げる「子」は相互に相手を観測して、たとえば強化学習のような方法を用いて接近/逃避行動を行なうことができる。このときは、他者の行動予測や対象に関する信念も不要である。

一方、障害物がある場合は、互いが遮蔽により見えなくなり、接近/逃避する対象の観測情報がなくなり、そのままでは行動決定ができない。ここで行動主体が対象についての信念(ex.位置、行動戦略など)を持つなら、最後に観測した地点からの相手の移動を予測して行動を継続できる。すなわち、障害物のある追跡課題には信念が必要である。

このような信念に関わる処理過程のモデルの検証のため、本研究では単純なゲーム世界を設定した(図 2)。この世界は 18×18 のグリッドワールドで、中央に 11×11 の四角い障害物があることで、観測の遮蔽を実現する。ここでオニと子は相互に他者を観測し、その動きを予想し、自己の行動を決定する。この世界では、オニが子を捕まえる戦略は2つある。一つはオニが子に向かって追跡行動をして、近づいて追いつめて捕まえるという戦略である。もう一つは、子が見えない間も逃げ続けると想定して、障害物を逆回りして子を待ちうけるという戦略である。いずれも、相手が見えない場合には自身の信念をもって行動する必要があり、さらにオニは子の内部状態を推定することで長期的な予想を行い、逆回りという戦略も検討できる。

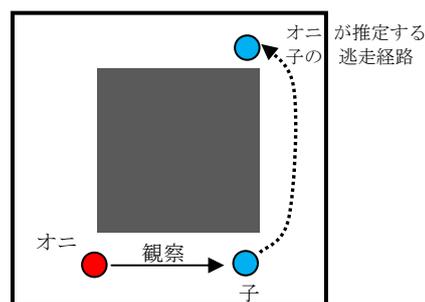


図2 「おにごっこ」ゲームで想定したゲーム世界

## 3. 追跡ゲームにおける信念推定モデル

### 3.1 自己の行動決定過程のモデル

このタスクでは、オニは子を追跡する戦略をとり、子はオニから逃げる戦略をとるのが基本である。これを基本戦略と呼ぶ。基本戦略では、相互の位置が見える場合には、オニは子の方向に移動し、子はオニから逃げる方向に移動する。その行動決定の方法として我々は、強化学習による行動決定を使用した。オニも子も事前に障害物のない世界で TD 学習による学習課題

を事前に行い、観測された相手の位置に応じて最適と思われる行動を獲得した。

相手が見えなくなった場合も、観測できない相手の位置を自身が信念としてもつ推定された相手に位置に置き換えること以外は、同様である。

TD 学習では、行動決定の際に期待報酬値の時間差を用いて確率的に行動を決定する。その際、期待報酬の差を正規化する意味で揺らぎの量を規定する温度  $T$  が使用される。温度が高い場合は探索的な揺らぎが大きい行動決定となり、温度が低い場合はわずかな報酬差を敏感に追従する行動決定となる。

温度のこの特性は、一貫性が実現する特性と一致する。一貫性が高い(温度が低い)ときは、行動主体は揺らぎの少ない直線的な行動で目標に向かい、一貫性が低い(温度が高い)ときは逆にノイズ、すなわちゆらぎの多い行動でゆっくりと追跡/逃走すると説明できる。

オニが子の長期間の行動予測を行うことで「障害物を逆方向に回って待ち受ける」という行動が可能となる。それを実現するには、他者が見えなくとも比較的長い時間でその動きを予測する能力と、ゲーム世界の任意の場所に自分が到達するまでの経路を探索するパスプランニングの能力が必要とされる。他者の動きと自己の経路の探索が十分に長時間に及んだとき、オニは障害物を逆回りする方がより効果的に子を捕まえられると判断できるようになる。

### 3.2 他者の行動の一貫性の評価

我々が想定する自己/他者の行動決定モデルでは、他者が現在の行動戦略に基づく動きを続けることを前提に、強化学習による行動予測を行うことができる。しかし前述のように、その行動決定にはゆらぎを表す温度/一貫性パラメータが含まれており、追跡/逃走の特性はそれにより大きく変化する。そのため、行動予測にはその一貫性パラメータの推定が欠かせない。

それに対して我々は、観測した他者の行動から、その揺らぎの程度を推定するモデルを構築し、計算機シミュレーションでその機能を確認した。本研究では、奥谷らによる行動の方向と目標①の方向のベクトルの内積により各瞬間の動きの妥当性を求め、Leaky Integrator で積分する方法で[奥谷 2012]、一貫性パラメータの推定を試みた。図 3 は他者のパラメータが、初めに 0.9、900 ステップで 0.1、1100 ステップで 0.9 となった場合の推定の過程で、100 ステップ程度で実用上困らない程度の推定に成功している。実用上とは、他者の行動のある範囲で正しく予想できるという意味で、これは次節で述べる他者の行動の誘導に大きく影響する。

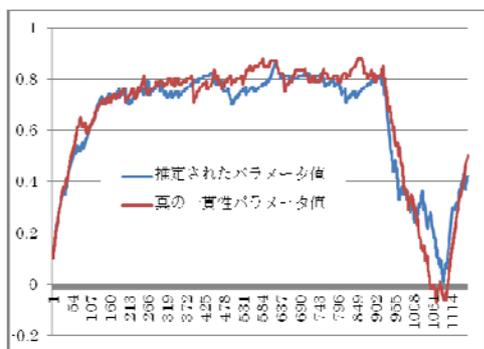


図 3 他者の一貫性パラメータの推定

### 3.3 だまし：相手の自己に対する推定の誘導

一貫性パラメータは他者の観測・推定され、結果として他者から見た自己の行動予測、ひいては他者の行動に影響する。であれば、たとえば他者が一時的に非常に一貫性の高い行動をとるなどして意図して行動の特性を変えたなら、それを推定する側はその一時的な行動特性をその相手の基本的な特性であると誤った推定をしてしまう。本研究では、一貫性の推定のこの特徴を利用して、他者をだます可能性について検討する。

本研究の「おにごっこ」では、相手が見えている間にその特性や信念を推定し、見えなくなった後はそれまでの特性や信念が継続すると仮定している。したがって、子が障害物に隠れる少し前にオニが高い一貫性をもって子を追跡すると、子はオニがその後もずっと継続して追跡してくると予測する。そこで、オニが「子はそのまま逃げ続ける」と長い予測をすると、子の予想される位置に到達する最短な経路が、子とは逆方向に障害物を回る戦略となるタイミングが発生する。

図 4 はその過程のシミュレーションの結果である。赤はオニ、青は子の移動の軌跡を表しており、細いブロックの高さが最初の瞬間からの予測の時間の経過に比例している。細く一番低いブロックが最初に予測した位置であり、だんだん高くなる方向に時間が経過して予測されている。太いブロックは、ある瞬間のオニと子の位置を示しており、赤と青で縁取りされた大きいブロックは、他者が推定したオニと子の位置である。この事例では、オニ(赤)は青で示される子が最初の位置からフィールドを左周りに逃げると予測し、実際に子(太い青)はそのように移動している。一方で子は、オニがすなおに追跡してくると信じてその①を推定しているが(左下の赤い縁取りの太いブロック)、オニは実際には逆回りの戦略を選択して逆方向に移動している(上側の太い赤いブロック)。結果として、子はオニが想定したように高い一貫性をもって追ってくると思っているオニから逃走し、オニは逆回りの経路でそれを待ち伏せることに成功している。

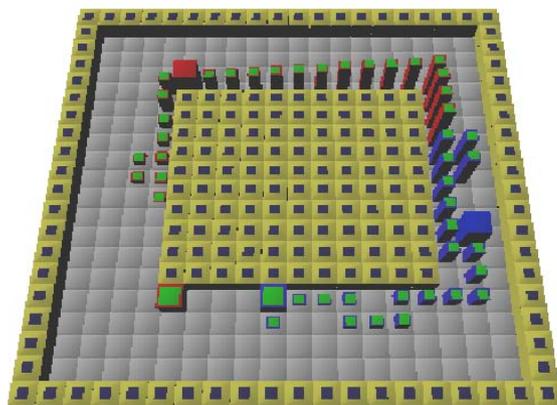


図 4 オニが、子のオニに対する一貫性推定を誘導することによる、「だまし」の実現。子がまっすぐ逃げる戦略を持たせ、オニは逆回りで子を捕まえようとしている。

## 4. 考察

### 4.1 信念の効用

一貫性パラメータの推定値を持つことは、対象の行動からその一貫性に関わる信念を読み取ることになる。その機能の利用は従来よりも強力な行動予測つながらと期待される。例えば「だまし」は、相手の内的状態を誘導して通常は取らない行動戦略をとらせ、その後にはその想定とは異なる行動をこちらが取ること

で実現される。今回の課題では、例えばオニが自己の一貫性を高く見せかけるためにしばらく高い一貫性で行動を続け、子のオニについての一貫性の推定値が十分大きくなるように誘導した。そして障害物で互いが見えなくなった瞬間、オニは長期のプランニングを行なった。この瞬間、子はオニの一貫性は高いと信じているため現在の逃走行動を長時間続けると予測された。そこでオニは子による自分についての行動予測を裏切り、障害物の逆側から回り込んで待ち伏せする行動をとれた。このように、他者の内的状態の過程についてのモデルを持つことは、単に相手の内部状態を推定して行動予測の精度を上げるだけでなく、より積極的な効用が見いだされる。

今回は、オニが子の行動の長期の予測を行い、経路探索により予想される子の位置への最適経路を発見するという行動を作りこんだ。このようなだまし行動が実現できるためには、本来的にはオニが子を「走らせ」たり「待ち伏せ」たりといった戦略についての価値を認識し、それらの間の価値を比較することで最適と思われる戦略を選択できる仕掛けが必要であろう。

「待ち伏せ」戦略の効用は、自己が追跡するという労力をかけずに相手に近づくことができるという、体力温存や省エネルギーの側面が強いように思われる。その価値を評価する機構を組み込んだとき、オニは単純な追跡以外の戦略を探索できる基盤ができると考えられる。

たとえば行動の「エネルギー」を考える可能性がある。行動によりエネルギーを消費すれば動けなくなる条件を導入すると、エネルギーを温存することに価値が与えられ、不必要な行動が削減される戦略が取られるであろう。従来、このような価値は課題を作る設計者が意図的にエージェントに組み込んでいる。しかしこのような信念を含む人間のモデル的理解をすすめるには、なぜ人間がそのような行動に価値を置くのか、ということも含めて原理を検討する必要があるであろう。

## 5. まとめ

行動決定における信念の機能を明らかにするため、信念の機能がなければ適切に解けないゲーム課題を用いし、それを解くためのエージェントモデルを提案した。信念を導入することで、エージェントは他者の信念、あるいは他者の自己に対する信念までも推定することで自身の行動決定をよりの確なものにできた。さらに本研究では「だまし」のような相手の判断の誘導までできる可能性を示した。「だまし」が必ずしも対人相互作用に必要という訳ではないが、相手の信念のダイナミクスを考えて相手の判断を誘導する程度の処理ができるなら、高いレベルの対人機能を持つ知的インタラクションを実現することにつながるであろう。今後は、本研究の成果を踏まえた上位の行動決定システムの検討を行っていく予定である。

## 参考文献

- [大沢 1989] 大沢：対話行為における知的行為者の計算モデル, コンピュータソフトウェア, Vol.6, No.4, pp348-359, (1989)
- [菅野 2006] 菅野太郎, 古田一雄：チーム協調モデルに基づく HA インタラクション戦略 -自然な人間-エージェント協調へ向けて-, HAI2006, 1-2A-2, 2006
- [横山 2009] 横山絢美, 大森隆司：協調課題における意図推定に基づく行動決定過程のモデル的解析, 電子情報通信学会論文誌 A, Vol.J92-A, No.11, pp.734-742, 2009
- [阿部 2011] 阿部香澄, 岩崎安希子, 中村友昭, 長井隆行, 横山絢美, 下斗米貴之, 岡田浩之, 大森隆司：子供と

遊ぶロボット：他者の状態推定に基づく行動決定モデルの適用, HAI2011, 1-2B-4, 2011

- [奥谷 2012] 奥谷一陽, 大森隆司：他者の信念推定に基づく行動決定のモデル化に向けた追跡ゲームの開発, HAI2012, 2D-17, 2012